Decoding Algorithm for Convolutional Codes under Linear Systems Point of View.

M. I. GARCÍA-PLANAS
Universitat Politècnica de Catalunya
Departament de Matemàtica Aplicada
Mineria 1, 08038 Barcelona
SPAIN
maria.isabel.garcia@upc.edu

El M. SOUIDI
Université Mohammed V
Département d'Informatique
B. P. 1014, Rabat
MOROCCO
emsouidi@yahoo.com

L.E. UM
Université Mohammed V
Département d'Informatique
B. P. 1014 , Rabat
MOROCCO
laurainlee@yahoo.fr

Abstract: In this work the output-observability of convolutional codes is examined and a decoding algorithm for convolutional codes using linear algebra techniques is presented.

Key–Words: Codes, linear systems, output-observability, generalized inverses.

1 Introduction

The convolutional codes were introduced by Elias [1] which suggested the use of a polynomial matrix $G(z)$ in the encoding procedure. These kind of codes are used extensively in numerous applications as satellite communication, mobile communication, digital video, radio among others.

One of the problems in convolutional codes theory was to find a method for constructing codes of a given rate and complexity with good free distance. Diverse methods have been introduced for this task.

There is a considerable amount of literature on the theory of convolutional codes over finite fields, (see [1, 2, 6, 8, 9, 11, 14, 17] for example).

A description of convolutional codes can be provided by a time-invariant discrete linear system called discrete-time state-space system in control theory (see [14, 16, 17]). We want to note that linear systems theory is quite general and it permits all kinds of time axes and signal spaces.

Rosenthal in [15], presented a first step toward an algebraic decoding algorithm for convolutional codes theory. It is based on an input/state/output description of the code and relies on the controllability matrix being the parity check matrix of an algebraically decodable block code.

The aim of this paper is to present an easy decoding algorithm using linear Algebra techniques.

2 Preliminaries

In this section, we present some basic notions about codes theory.

Definition 1 *An error correcting code $\mathcal{C} \subseteq \mathcal{A}^*$ is said that is a convolutional code, when $\mathcal{C}$ is linear (considered as a vector space over $\mathbb{F}_q = GF(q)$ (the Galois field of q elements) with the usual sum of vectors) invariant by right translation operator and has compact support.*

Following Rosenthal and York [16], a convolutional code is defined as a submodule of $\mathbb{F}^n[s]$.

Definition 2 *Let $\mathcal{C} \subseteq \mathcal{A}^*$ be a code. Then $\mathcal{C}$ is a convolutional code if and only if $\mathcal{C}$ is a $\mathbb{F}[s]$-submodule of $\mathbb{F}^n[s]$.*

Corollary 3 *There exists an injective morphism of modules*

$$\begin{aligned} \psi : \mathbb{F}^k[s] &\longrightarrow \mathbb{F}^n[s] \\ u(s) &\longrightarrow v(s). \end{aligned}$$

Equivalently, there exists a polynomial matrix $G(s)$ (called encoder) of order $n \times k$ and having maximal rank such that

$$\mathcal{C} = \{v(s) \mid \exists u(s) \in \mathbb{F}^k[s] : v(s) = G(s)u(s)\}.$$

The rate k/n is known as the ratio of a convolutional code. We denote by ν_i the maximum of all degrees of each of the polynomials of each line, we define the complexity of the encoder as $\delta = \sum_{i=1}^{n} \nu_i$, and finally we define the complexity convolution code $\delta(\mathcal{C})$ as the maximum of all degrees of the largest minors of $G(s)$.

The representation of a code by means a polynomial matrix is not unique, but we have the following proposition.

Proposition 4 *Two $n \times k$ rational encoders $G_1(s)$, $G_2(s)$ define the same convolutional code, if and only if there is a $k \times k$ unimodular matrix $U(s)$ such that $G_1(s)U(s) = G_2(s)$.*

After a suitable permutation of the rows, we can assume that the generator matrix $G(s)$ is of the form

$$G(s) = \begin{pmatrix} P(s) \\ Q(s) \end{pmatrix}$$

with right coprime polynomial factors $P(s) \in \mathbb{F}^{(n-k)\times k}$ and $Q(s) \in \mathbb{F}^{k\times k}$, respectively.

2.1 Convolutional code as input-state-output

This section is based on results obtained by the authors in [4].

Consider the matrices $A \in \mathbb{F}^{\delta\times\delta}$, $B \in \mathbb{F}^{\delta\times k}$, $C \in \mathbb{F}^{(n-k)\times\delta}$ and $D \in \mathbb{F}^{(n-k)\times k}$. A convolutional code $\mathcal{C}$ of rate k/n and complexity δ can be described by the following linear system of equations:

$$\left.\begin{array}{rl} x_{t+1} & = Ax_t + Bu_t \\ y_t & = Cx_t + Du_t \end{array}\right\}, \quad \begin{array}{rl} v_t & = \begin{pmatrix} y_t \\ u_t \end{pmatrix}, \\ x_0 & = 0. \end{array} \tag{1}$$

In terms of systems theory the variable x_t is called a state variable of the system at time t, u_t the input vector and y_t the vector output obtained from the combination of input and state variable. If no confusion is possible, we will write the system as the quadruple of matrices (A, B, C, D).

In terms of the theory of codes, we have the input of the encoder after time t which is called the information o vector message u_t; the vector y_t created by the encoder is called parity vector, the code vector v_t is transmitted via the communication channel. We will denote the code convolution created in this way, by $\mathcal{C}(A, B, C, D)$.

The transfer matrix

$$G(s) = C(sI - A)^{-1}B + D$$

provides a rational description of the convolutional code.

In terms of the input-state-output representation of a convolutional code, the free distance of a convolutional code $\mathcal{C}$, that is, the minimum Hamming distances between any two code sequences of $\mathcal{C}$, can be characterized as (see [7])

$$d_{free}(\mathcal{C}) = \lim_{j\to\infty} d_j^c(\mathcal{C}), \tag{2}$$

where

$$d_j^c(\mathcal{C}) = \min_{u(0)\neq 0} \left\{ \sum_{t=0}^{j} wt(u(t)) + \sum_{t=0}^{j} wt(y(t)) \right\}$$

is the j-th column distance of the convolutional code $\mathcal{C}$, for $j = 0, 1, 2, \ldots$.

The free distance of a convolutional code determines to a large extend the error rate in the case of maximum likelihood decoding, and is a good indicator of the error correcting performance of the code.

We will note that the concept of minimality of an input-state-output representation is different from the concept of minimality of a representation, in classical linear systems theory. A representation (A, B, C, D) in linear systems literature is minimal if and only if the pair (A, B) is controllable and the pair (A, C) is observable. In fact, if the pair (A, B) is controllable, then the observability of the pair (A, C) ensures that the linear system 1 describes a noncatastrophic convolutional encoder, as we can see in the following lemma.

Lemma 5 (Lemma 2.11 of [16],) *Assume that the pair of matrices (A, B) is controllable. The convolutional code $C(A, B, C, D)$ defined through 1 represents an observable convolutional code if and only if the pair of matrices (A, C) is observable.*

Remember that a convolutional code is said catastrophic if it prone to catastrophic error propagation, i.e. a code in which a finite number of channel errors causes an infinite number of decoder errors. Any given convolutional code is or is not a catastrophic code.

Related to the minimality realization of an encoder is the output-observability property.

Output-observability represents the possibility of an internal state, to be only defined by a finite set of outputs, for a finite number of steps.

Definition 6 *The system (A, B, C, D) is said be output observable if the state sequence $x(0), \ldots, x(\ell)$ is uniquely determined by the knowledge of the output sequence $y(0), \ldots, y(\ell)$ for a finite number of steps $\ell \in \mathbb{N}$.*

Observe that $x(1), \ldots, x(\ell)$ are determined by the knowledge of $x(0)$ and $u(0), \ldots, u(\ell-1)$ because

of

$$\begin{aligned} x(1) &= Ax(0) + Bu(0) \\ x(2) &= Ax(1) + Bu(1) = \\ &= A^2x(0) + ABu(0) + Bu(1) \\ &\vdots \\ x(\ell) &= Ax(\ell-1) + Bu(\ell-1) = \\ &= A^\ell x(0) + A^{\ell-1}Bu(0) + \ldots + \\ &\quad + ABu(\ell-2) + Bu(\ell-1), \end{aligned}$$

and the elements $x(0)$, and $u(0), \ldots, u(\ell)$ can be obtained solving the following system of matrix equations.

$$\begin{aligned} y(0) &= Cx(0) + Du(0) \\ y(1) &= Cx(1) + Du(1) = \\ &= CAx(0) + CBu(0) + Du(1) \\ &\vdots \\ y(\ell) &= Cx(\ell) + Du(\ell) = \\ &= CA^\ell x(0) + CA^{\ell-1}Bu(0) + \ldots + \\ &\quad + CBu(\ell-1) + Du(\ell) \end{aligned} \tag{3}$$

In a more general way we can define the output-observability character saying that the state sequence $x(s), \ldots, x(\ell)$ is uniquely determined by the knowledge of the output sequence $y(s), \ldots, y(s+\ell)$ for a finite number of steps $\ell \in \mathbb{N}$.

In an analogous way we have that $x(s+1), \ldots, x(s+\ell)$ are determined by the knowledge of $x(s)$ and $u(s), \ldots, u(s+\ell-1)$ because of

$$\begin{aligned} x(s+1) &= Ax(s) + Bu(s) \\ x(s+2) &= Ax(s+1) + Bu(s+1) = \\ &= A^2x(s) + ABu(s) + Bu(s+1) \\ &\vdots \\ x(s+\ell) &= Ax(s+\ell-1) + Bu(s+\ell-1) = \\ &= A^{s+\ell}x(s) + A^{s+\ell-1}Bu(s) + \ldots + \\ &\quad + ABu(s+\ell-2) + Bu(s+\ell-1), \end{aligned}$$

and the elements $x(s)$, and $u(s), \ldots, u(s+\ell)$ can be obtained solving the following system of matrix equations.

$$\begin{aligned} y(s) &= Cx(s) + Du(s) \\ y(s+1) &= Cx(s+1) + Du(s+1) = \\ &= CAx(s) + CBu(s) + Du(s+1) \\ &\vdots \\ y(s+\ell) &= Cx(s+\ell) + Du(s+\ell) = \\ &= CA^{s+\ell}x(s) + CA^{s+\ell-1}Bu(s) + \ldots + \\ &\quad + CBu(s+\ell-1) + Du(s+\ell) \end{aligned} \tag{4}$$

Calling $T_\ell(A, B, C, D)$ (that we simply write T_ℓ if confusion is not possible) the matrix

$$T_\ell = \begin{pmatrix} C & D & & & & \\ CA & CB & D & & & \\ CA^2 & CAB & CB & D & & \\ \vdots & & & \ddots & \ddots & \\ CA^\ell & CA^{\ell-1}B & CA^{\ell-2}B & \ldots & CB & D \end{pmatrix} \tag{5}$$

We have the following.

Proposition 7 *A system (A, B, C, D) is output observable if and only if the matrix T_ℓ has full row rank for all $\ell \in \mathbb{N}$.*

Proof:

First of all, we observe that for each ℓ, the matrix T_ℓ is the corresponding matrix to the system (3). So, if the number of rows is bigger than the number of columns, there are values of $y(0), \ldots, y(\ell)$, for which the system has no solution. □

Corollary 8 *A necessary condition for output-observability of the system (A, B, C, D) is that the matrix $\begin{pmatrix} C & D \end{pmatrix}$ has full row rank*

Therefore, we assume that the number of rows is less than or equal to the number of columns. It is well know that in this case and for each ℓ, the systems (3) have solution for all $y(0), \ldots, y(\ell)$, if and only if the systems have full rank.

Corollary 9 *If the matrix D in the system (A, B, C, D) has full row rank, the system is output-observable.*

3 Decoding problem

It is well known the existence of several algorithms for the decoding of convolutional codes. Foremost among them, codes are decoded using the so called Viterbi decoding algorithm. The Viterbi Algorithm was first proposed as a solution to the decoding of convolutional codes by Andrew J. Viterbi in 1967, [18],

In order to analyze this process we will assume that a certain code word $\{v_t\}_{t\geq 0} = \{(y)_t u_t\}$ was sent and the message word $\{\widehat{v}_t\}_{t\geq 0} = \left\{ \begin{pmatrix} \widehat{y}_t \\ \widehat{u}_t \end{pmatrix} \right\}$ been received. The decoding problem then asks for the minimization of the error

$$\begin{aligned} \text{error} &= \min_{\{v_t\}\in C} \sum_{t=0}^{\infty} \text{dist}\,(v_t, \widehat{v}_t) \\ &= \min\,(\textstyle\sum_{t=0}^{\infty}(\text{dist}\,(y_t, \widehat{y}_t) + \text{dist}\,(u_t, \widehat{u}_t)) \end{aligned} \tag{6}$$

If in the transmission, no errors are produced, then $\{\widehat{v}_t\}_{t\geq 0}$ is a valid trajectory and the error value defined in 6 is zero.
Otherwise, if the error value isn't null, then the sequence received isn't a codeword, and doesn't belong to the code family. Then, comes the importance of decoding, which consists of finding out, from the gotten sequence the encoded word supposed to have been received.

3.1 Decoding convolutional codes

We are interested in the decoding of convolutional codes represented as linear systems.

Using the matrix 5 we obtain a representation in terms of state input-output of the code

$$\begin{pmatrix} C & D & & & & \\ CA & CB & D & & & \\ CA^2 & CAB & CB & D & & \\ \vdots & & & \ddots & \ddots & \\ CA^{\ell} & CA^{\ell-1}B & CA^{\ell-2}B & \dots & CB & D \end{pmatrix} \begin{pmatrix} x(0) \\ u(0) \\ \vdots \\ u(\ell) \end{pmatrix} = \begin{pmatrix} y(0) \\ y(1) \\ \vdots \\ y(\ell) \end{pmatrix} \tag{7}$$

Proposition 10 *Let (A, B, C, D) a representation of an output-observable code. Then, the system 7 is solvable.*

Proof: If the system (A, B, C, D) is output-observable the matrix of the equation 7 has full row rank. □

Remark 11 *It is usual to consider the initial state of the system $x(0) = 0$. In this case the system 7 is reduced to*

$$\begin{pmatrix} D & & & & \\ CB & D & & & \\ CAB & CB & D & & \\ \vdots & & \ddots & \ddots & \\ CA^{\ell-1}B & CA^{\ell-2}B & \dots & CB & D \end{pmatrix} \begin{pmatrix} u(0) \\ \vdots \\ u(\ell) \end{pmatrix} = \begin{pmatrix} y(0) \\ y(1) \\ \vdots \\ y(\ell) \end{pmatrix}. \tag{8}$$

So, in this case the solvability of the system is ensured if the matrix

$$\widehat{T}_{\ell-1} = \begin{pmatrix} D & & & & \\ CB & D & & & \\ CAB & CB & D & & \\ \vdots & & \ddots & \ddots & \\ CA^{\ell-1}B & CA^{\ell-2}B & \dots & CB & D \end{pmatrix}$$

has full rank.

But, if the matrix of the system 7 has full row rank, the system 8 is not necessarily solvable as we can see in the following example

Example 1. Let (A, B, C, D) a realization of a convolutional code with $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 1 \end{pmatrix}$ and $D = (0)$, the system

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \\ u(0) \\ u(1) \end{pmatrix} = \begin{pmatrix} y(0) \\ y(1) \end{pmatrix}$$

is compatible for all $\begin{pmatrix} y(0) \\ y(1) \end{pmatrix}$ and the solution is $x_1 = y(1)$, $x_2 = y(0)$, nevertheless the system

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \end{pmatrix} = \begin{pmatrix} y(0) \\ y(1) \end{pmatrix}$$

has only solution for $y(0) = y(1) = 0$. That is to say the initial condition for the system are restrictive conditions for solving the system.

But, in any case, we have the following proposition.

Proposition 12 *If the matrix $\widehat{T}_{\ell-1}$ has full row rank the system 7 is solvable with initial conditions $x(0) = 0$.*

Proof: If the matrix $\widehat{T}_{\ell-1}$ has full row rank, the system 8 is solvable. Then, if $(u(0), \dots, u(\ell))$ is a solution of this system, clearly $(0, u(0), \dots, u(\ell))$ is a solution for the system 7. □

3.2 Solving 7

In the case where the matrix of the system 7 doesn't have full row rank, the existence of the solution is not guaranteed and depends on $\begin{pmatrix} y(0) \\ y(1) \end{pmatrix}$.

If the system is not compatible we can find approximate solutions using generalized inverses matrices as the Moore-Penrose pseudoinverse matrix.

Remember that, given a matrix $A \in M_{n\times m}(\mathbb{F})$ a matrix $X \in M_{m\times n}(\mathbb{F})$ is called generalized inverse if and only if it verifies

a) $AXA = A$,

A generalized inverse X of A is called a reflexive generalized inverse if and only if it verifies

b) $XAX = X$

A reflexive generalized inverse X of A is called normalized and will be denoted by A^{nor} if and only if it verifies

c) $(AA^{nor})^t = AA^{nor}$

and finally a normalized generalized inverse A^{nor} is called the Moore-Penrose pseudoinverse and will be denoted by A^{+} if and only if verifies

d) $(A^+A)^t = A^+A$.

A linear system $Ax = y$ can be solved if we have a generalized inverse of the matrix A

observe that if

$$Ax = y$$

we have

$$AXAx = y$$

so

$$AXy = y$$

that is to say

$$Xy$$

is a solution and the solution general can be easily obtained if we are taking into account that Im $(I_m - XA) =$ Ker A.

Not always there exists the normalized and pseudoinverse matrix. Penrose [13] showed that every matrix A over the complex field has a normalized inverse and a unique A^+. However, Pearl [12] showed that a matrix $A \in M_{m\times n}(\mathbb{F})$ of rank r over an arbitrary field has a normalized and a Moore-Penrose A^+ (unique) only under certain conditions. In fact we have the following result

Theorem 13 *Let A be an $m \times n$ matrix of rank r over a field F. Then, A has a normalized generalized inverse A^{nor} if and only if*

$$r = \text{rank}\,(A^tA).$$

And A has a Moore-Penrose pseudoinverse A^+ if and only if

$$r = \text{rank}\,(A^tA) = \text{rank}\,(AA^t).$$

Example 2. Over $\mathbb{F}_5$ the 1×5-matrix $A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$ doesn't have a A^+ pseudoinverse.

For those methods of solving, instead of only detect the error, (as for instance in the Viterbi decoding algorithm), and put out the correct sequence that should have been received, at the same time we detect the error, and give the original message before encoding. In order to solve the system, the algorithms we will be using will be divided between some steps.

3.3 Iterative decoding algorithm

The first algorithm for solving focuses more on directly correcting the error in case of disturbance by approaching the original input that is supposed to have been encoded (and whose encoding was disturbed, in case of disturbance). We will consider that if we get as close as possible to the initial input, with the output received, by approaching the best possible solution of our system, then we will have the error implicitly corrected and the original message. We will do so considering initial conditions, and output-observability matrix.

Of course initial conditions are not our main concern.

Suppose now, that there exists D^+ or any other generalized inverse X

Iterative process

Step 1 Look the number if outputs, denoted by ℓ.

Step 2

Case 1 Fix $x(0)$ such that $x(0) = 0$, then solve $Du(0) = y(0)$:

$D^+y(0)$ (or $Xy(0)$) is an approximation solution (that is exact if $Du(0) = y(0)$ is compatible, in particular if D has full row rank).

Case 2 Fix $x(0)$ such that $x(0) \neq 0$ then solve $(C\ D)\begin{pmatrix} x(0) \\ u(0) \end{pmatrix} = y(0)$:

$(C\ D)^+(y(0))$ (or $Xy(0)$) is an approximation solution (that is exact if $(C\ D)\begin{pmatrix} x(0) \\ u(0) \end{pmatrix} = y(0)$ is compatible, in particular if $(C\ D)$ has full row rank).

In both cases the error of solution is minimized consodering $d_H(T_0\begin{pmatrix} x(0) \\ u(0) \end{pmatrix}, y(0))$ and u's Hamming weight as well. Then, settle for the approximate minimal solution.

Step 3: Iteratively, solve

$Du(\ell) = y(\ell) - (CA^\ell x(0) + CA^{\ell-1}Bu(0) + \ldots + CBu(\ell-1))$:

$D^+(y(\ell) - (CA^\ell x(0) + CA^{\ell-1}Bu(0) + \ldots + CBu(\ell-1)))$ (or $X(y(\ell) - (CA^\ell x(0) + CA^{\ell-1}Bu(0) + \ldots + CBu(\ell-1))))$ is an approximate solution.

Example 3. Let's look at the decoding of words, with the code in which $p \geq k$ defined as: in the field $\mathbb{F}_7$, let (A_1, B_1, C_1, D_1) with

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 1 \end{pmatrix},\ B = \begin{pmatrix} 0 \\ 2 \end{pmatrix},\ C = \begin{pmatrix} 5 & 2 \\ 0 & 6 \\ 3 & 0 \end{pmatrix},\ D = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix} \tag{9}$$

Let's try to decode the sequence: $y = (y(0), y(1), y(2)) = (3\,1\,5,\ 6\,0\,2,\ 4\,2\,2)$

Observe that in this case, the system is not output-observable. Step 1: We have $\ell = 2$

Step 2: We have

$$D = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix};$$

then we solve $\begin{pmatrix} C & D \end{pmatrix} \begin{pmatrix} x(0) \\ u(0) \end{pmatrix} = \bar{y}(0)$.

We have: rank $D = 1$, which means that D has full (column) rank; we decide to fix $x(0) = 0$

So we solve $\begin{pmatrix} D \end{pmatrix} \begin{pmatrix} u(0) \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix} \begin{pmatrix} u(0) \end{pmatrix} = \bar{y}(0) = \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix}$.

The system is clearly incompatible, so we try to solving using a pseudoinverse matrix.

The matrix D verifies conditions for existence of D^+.

In this particular case the pseudoinverse of such a matrix D is given by: $D^+ = (D^t D)^{-1} D^t$. So, we get: $D^+ = \begin{pmatrix} 1 & 6 & 3 \end{pmatrix}$ Then,

$$\begin{pmatrix} 1 & 6 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 3 \end{pmatrix} = u(0)$$

We get: $u(0) = 3$; when verifying: $\begin{pmatrix} D \end{pmatrix} \begin{pmatrix} u(0) \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ 4 \end{pmatrix}$ Here, at least we detect errors on 2 elements of the sequence. Indeed, we get for $x(0) = (0)$, $u(0) = (3)$.

Step 3: Solve $\widehat{T}_1 \begin{pmatrix} u(0) \\ u(1) \end{pmatrix} = \begin{pmatrix} \bar{y}(0) \\ \bar{y}(1) \end{pmatrix}$ with $u(0) = 3$. So, it suffices to solve $\begin{pmatrix} CB & D \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \end{pmatrix} = \bar{y}(1) = \begin{pmatrix} 6 \\ 0 \\ 2 \end{pmatrix}$

Then, we solve

$$\begin{pmatrix} 4 & 2 \\ 5 & 5 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ u(1) \end{pmatrix} = \bar{y}(1) = \begin{pmatrix} 6 \\ 0 \\ 2 \end{pmatrix}$$

So, $\begin{pmatrix} D \end{pmatrix} \begin{pmatrix} u(1) \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 2 \end{pmatrix} - 3 \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \\ 2 \end{pmatrix}$ which means: $D^+ \begin{pmatrix} 1 \\ 6 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 6 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 6 \\ 2 \end{pmatrix} = 1 = u(1)$. Then the solution $u(1)$ is (1). When we verify, $\begin{pmatrix} CB & D \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 6 \end{pmatrix} \neq \begin{pmatrix} 6 \\ 0 \\ 2 \end{pmatrix}$. So, we detected errors in the second sequence, and our approximate solution $u(1)$ is 1.

Finally, we solve $\widehat{T}_2 \begin{pmatrix} u(0) \\ u(1) \\ u(2) \end{pmatrix}$ with $u(0) = 3$, and $u(1) = 1$. So we solve $\begin{pmatrix} CAB & CB & D \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \\ u(2) \end{pmatrix} = \bar{y}(2) = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}$ with $u(0) = 3$, and $u(1) = 1$.

Then, we solve

$$\begin{pmatrix} 6 & 4 & 2 \\ 5 & 5 & 5 \\ 4 & 0 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ u(2) \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}$$

We get: $Du(2) = \begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix}$; we already have $D^+ = \begin{pmatrix} 1 & 6 & 3 \end{pmatrix}$.

Which means that: $D^+ \begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix} = u(2)$. The approximate solution $u(2)$ is (5). Verifying, we have:

$$\begin{pmatrix} CAB & CB & D \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$$

However, $d_H((4,\,3,\,0),(4,\,2,\,2)) = 2$; so, we detected 2 errors, and we approached the solution the best way possible.

For this case, $(\bar{y}(0), \bar{y}(1), \bar{y}(2)) = (3\,1\,5,\ 6\,0\,2,\ 4\,2\,2)$ there were multiple errors during transmission.

The decoded sequence is $u = (u(0), u(1), u(2)) = (3,\ 1,\ 5)$, with initial condition: $x(0) = (0\,0)$

Remark 14 *This method is quite efficient for error detection; indeed, we can tell when there was a mistake within a sequence, by computing: $d_H(y,\ \bar{y})$, y the output obtained from the approximate solution; however the correction rate is harder to figure out, since we only detect when a mistake occurs, and we assume the solution we get is the closest without any verification.*

3.4 Output-observability matrix and Syndrome former matrix

Let (A, B, C, D) be a realization of a convolutional code.

From the system 7, we can deduce the syndrome former matrix for the given code.

Proposition 15 *Suppose that $\ell \geq \delta$. By making elementary transformations to matrix equation 7 we can deduce the syndrome former matrix for the convolutional code.*

Proof: The system 7 can be rewritten as

$$\begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^\ell \end{pmatrix} x(s) = \begin{pmatrix} \widehat{T}_{\ell-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} -u(s) \\ -u(s+1) \\ \vdots \\ -u(s+\ell) \\ y(s) \\ y(s+1) \\ \vdots \\ y(s+\ell) \end{pmatrix} \tag{10}$$

Now, and taking into account that $\ell \geq \delta$ there exist an invertible matrix $P \in Gl(p \times \ell, \mathbb{F})$ such that

$$P \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^\ell \end{pmatrix} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{\delta-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \mathcal{O} \\ 0 \end{pmatrix}$$

where $\mathcal{O}$ is the observability matrix of the pair (A, C).

Applying the matrix P to the matrix equation 10 we obtain

$$\left(\begin{array}{c} \mathcal{O} \\ \hline 0 \end{array}\right)(x(s)) = \left(\begin{array}{c|c} M_1 & M_2 \\ \hline M_3 & M_4 \end{array}\right) \left(\begin{array}{c} -u(s) \\ \vdots \\ -u(s+\ell) \\ \hline y(s) \\ \vdots \\ y(s+\ell) \end{array}\right) \tag{11}$$

Then, $\left(\begin{array}{c|c} M_3 & M_4 \end{array}\right)$ is the syndrome former matrix. □

4 Conclusions

In this paper the output-observability of convolutional codes has been revised and a decoding algorithm for convolutional codes using linear algebra techniques has been presented. Also a syndrome former matrix from output-observability matrix has been deduced.

Acknowledgements: The first author is supported by grant MTM2010-19356-C02-02.

References:

[1] P. Elias, *Coding for Noisy Channels*, IRE Conv.Rec. **4**, pp. 37-46, (1955) .

[2] Ch. Fragouli, R. D. Wesel, Convolutional Codes and Matrix Control Theory, *Proceedings of the 7th International Conference on Advances in Communications and Control*, Athens, Greece, 1999.

[3] G. D. Forney, *Convolutional codes: Algebraic structure*. IEEE trans. Information Theory, 1970.

[4] Mª I. García-Planas, El M. Souidi, L.E. Um, Convolutional codes under linear systems point of view. Analysis of output-controllability, *Wseas Transactions on Mathematics*, 11, (4), 2012, pp. 324-333.

[5] Mª I. García-Planas, Bifurcation diagrams of generic families of singular systems under proportional and derivative feedback. *Wseas Transactions on Mathematics*, 7, (2), 2008, pp. 1–11.

[6] H. Gluesing-Luerssen., U. Helmke, J.I. Iglesias Curto Algebraic Decoding for Doubly Cyclic Convolutional Codes, Advances in Mathematics of Communications 4(2010), 83-99 with G. Schneider) State space realizations and monomial equivalence for convolutional codes , Linear Algebra and its Applications **425**, (2007), 518-533

[7] R. Hutchinson, J. Rosenthal, R. Smarandache, *Convolutional codes with maximum distance profile*, Systems Control Lett. 54 (1), pp- 53-63, (2005).

[8] J. I. Iglesias, *A study on convolutional Codes. Classification, new families and decoding*, Tesis Doctoral, (2007).

[9] M. Kuijper, R. Pinto, *On minimality of convolutional ring encoders*. IEEE Trans. on Information Theory, **55**, (11), pp. 4890-4897, (2009).

[10] H. Loeliger, G. D. Forney, T. Mittelholzer, M.D. Trot, *Minimality and observability of group systems*, Linear Algebra and its Applications, **205-206**, pp. 937-963, (1994).

[11] J. L. Massey, M. K. Sain, *Codes, Automata and continuous systems: explicit interconnections.* IEEE Trans. on Automatic Control, Vol AC-12 (6), pp.644-650, (1967).

[12] M.H. Pearl, Generalized inverses of matrices with entries taken from an arbitrary field, *Linear Algebra and Appl*, 1, 1968 pp. 571–587.

[13] R. Penrose, A generalized inverse of matrices, *Proc. Cambridge Philos. Sot*. 51, 1955, pp. 406–413 .

[14] J. Rosenthal, *Some interesting problems in systems theory wich are of fundamental importance in coding theory*, Proceedings of the 36th IEEE Conference on Decision and Control, (1997).

[15] J. Rosenthal, *An algebraic decoding algorithm for convolutional codes*. In G. Picci and D. S. Gilliam, editors, Dynamical Systems, Control, Coding, Computer Vision; New Trends, Interfaces, and Interplay, pages 343360. Birkhäuser, Basel, (1999).

[16] J. Rosenthal, E. V. York, *BCH Convolutional Codes*, IEEE Trans. Information Theory vol. 45 (6), 1833-1844, (1999).

[17] J. Rosenthal, J. M. Schumacher, E. V. York, *On Behaviors and Convolutional Codes*, IEEE Trans. on Information Theory, **42**, (6), pp. 1881-1891, (1996).

[18] A.J. Viterbi, *Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm*, IEEE Transactions on Information Theory, **13**, (2), (1967), pp. 260-269.